QUALITY CONTROL METHOD FOR MANUFACTURING BIOPOLYMER ARRAYS

Field of the Invention

[0001] The invention relates to a quality control method for manufacturing biopolymer arrays comprising the use of detectable protecting groups.

Background of the Invention

[0002] The synthesis of nucleic acids and peptides on a solid phase has been an established process for the last 20 years. The most prevalent method of nucleic acid synthesis is the phosphoramidite method of Beaucage and Caruthers, Tetrahedron Lett. 22 (1981), 1859-1862, where the oligonucleotide chain is built up by the repetitive condensation of individual nucleotide building blocks in the 3' or 5' direction. A variety of orthogonal protecting groups are used to protect three reactive nucleotide groups: the ribose sugar 5' hydroxyl group, the amino protecting group of the nucleobases adenine, guanine and cytosine (thymine does not need a protecting group), as well as the phosphate protecting group of the nucleotide 3' phosphate residue. These protecting groups are then cleaved off under varying conditions, either during or after the synthesis. The 4,4'-dimethoxytriphenylmethyl (DMT) group has become the standard for 5' hydroxyl, the 2-cyanoethyl group the standard for phosphate residue and various acyl groups the standard for the amino functions of the nucleobases according to Büchi and Khorana (J. Mol. Biol. 72 (1972), 251-258) and Souveaux (in: "Methods in Molecular Biology" Vol. 26, Chap. 1 Protocol for Oligonucleoside Conjugates, Ed. S. Agrawal, Humana Press Inc., Totowa, N.J. 1994). The DMT group is cleaved off during synthesis in order to generate an hydroxyl group to which the next phosphoramidite can bind. The other named protecting groups remain until the end of the synthesis in order to prevent any side-reactions or by-products. At the end of the synthesis the complete oligonucleotide is completely deprotected largely by means of a base, whereby 2-cyanoethyl and acyl protecting groups are cleaved off.

[0003] There are essentially two means of producing biochips: off-chip and on-chip synthesis of oligonucleotide probes. In off-chip synthesis, the oligonucleotide is produced on a commercially available synthesizer using the above-mentioned standard reagents and then immobilized on the chip. In on-chip synthesis, the oligonucleotide is produced directly on the chip using the above-mentioned standard reagents. In the former case, the quality of the oligonucleotide can be analyzed by means of analytical processes such as HPLC or mass spectrometry and, where necessary, improved via purification. The latter case of on-chip synthesis allows for only limited quality control, and purification is not possible. Quality control is normally only possible by means of the covalent binding of a (mainly fluorescent) coloring material at the terminus of the oligonucleotide, which can then be detected and quantified.

[0004] There is scarcely any mention in the literature of processes for determining the deprotection degree of oligonucleotides. A first method comprises the detection of nucleobases and 5'-hydroxyl protecting groups on oligonucleotides with monoclonal antibodies, Fu et al., Analytical Biochemistry (2002), 306(1), 135-143). Other references describe the use of cleavable, fluorescent protecting groups. These refer, however, to protecting groups for the 5'-hydroxyl group or the phosphate residue, but not for the nucleobase amino groups. In this context, it should be noted that US Patent 6,238,862 B1 describes a fluorescent, photo labile protecting group for the 5'-hydroxyl group used for determining the synthesis efficiency of the DNA synthesis array. Wagner and Pfleiderer (Helv. Chim. Act. 80 (1990), 200-212) describe a nucleobase protecting group which basically has fluorescent properties. However, this group has been developed for other purposes, especially for improved deprotection properties by using a β-elimination reaction process.

[0005] As a rule, exact quality controls cannot be carried out with on-chip synthesis. Presently, only the binding of dyes to the terminus of the oligonucleotide gives an indirect indication of the quantity of solid phase oligonucleotide and of the quality of the synthesis. This does not, however, provide any indication of the rate of the deprotection at the end of the synthesis, i.e., whether all 2-cyanoethyl and acyl protecting groups were cleaved off. It

is known that, under standard conditions, the nucleobase acyl protecting groups cannot always be quantitatively cleaved off, and part of the amino groups therefore remain protected. Complete deprotection of the amino groups is, however, essential for the optimal application of the chip. Should not all of the amino function protecting groups be cleaved off, a subsequent hybridization can be decisively impaired, since the formation of Watson-Crick or Hoogsteen base pairs will be inhibited.

Summary of the Invention

[0006] Thus, the invention relates to a quality control method for manufacturing biopolymer arrays comprising:

(a) synthesizing a plurality of different biopolymer species on an array from monomeric or oligomeric building blocks comprising detectable protecting groups;

(b) optionally carrying out a determination of the detectable protecting groups on the array after synthesis;

(c) cleaving off the detectable protecting groups, and

(d) carrying out a determination of the detectable protecting groups on the array after cleavage in order to determine the efficacy of deprotection.

[0007] These simple steps would enable the quality control of the synthesized biopolymers. The main focus is quality control of the completeness of the biopolymer deprotection. The key advantage is that the methodology is non-destructive and requires no further steps once the final deprotection or detection has been carried out. The outcome is an evaluation of the rate of deprotection as well as of the quality of freely accessible biopolymers for later use.

Brief Description of the Drawings

[0008] Further, the invention shall be explained in more detail by the following figures.

[0009] Figure 1: Fluorescent, pyrene-labeled cytidine building block. The pyrene-containing nucleobase protecting group exhibits fluorescent properties and can therefore be detected by means of optical analysis procedures.

[0010] If the nucleobase amino functions are protected with a cleavable acyl group carrying a biotin or digoxigenin labeled system, then detection can be carried out using streptavidin or anti-digoxigenin.

[0011] Figure 2: A cytidine building block with an IR-detectable nitro aromatic moiety.

[0012] Figure 3: A cytidine building block with an electrochemically detectable ferrocene moiety.

[0013] Figure 4: An adenine building block with a fluorescent detectable stilbene moiety.

[0014] Figure 5: Scheme of the synthesis of an adenine building block of Figure 4.

Detailed Description of the Invention

[0015] Generally, the invention relates to the use of detectable protecting groups in the manufacture of biopolymer arrays. The term "biopolymer" as used in the present application particularly relates to nucleic acids such as DNA or RNA or nucleic acid analogs such as peptide nucleic acids (PNA) or locked nucleic acids (LNA) or combinations thereof. The term, however, also relates to peptides and peptide analogs as well as to other biopolymers such as carbohydrates or any combinations thereof. Preferably the biopolymer species are selected from nucleic acids and nucleic acid analogs wherein the detectable protecting groups are coupled to nucleobases, particularly to amino groups of nucleobases.

[0016] The term "array" as used in the present application relates to a solid support, e.g., a planar or non-planar solid support. The solid support may have a surface selected from metals such as silicon, metal oxides such as silica, glass or plastic or any combination thereof. Preferably the array is a chip.

[0017] The method of the present invention comprises synthesizing a plurality of different biopolymer species on the array from monomeric or oligomeric building blocks. For example, nucleic acids may be synthesized from phosphoramidite or phosphonate building blocks as known in the art. The synthesis is a spatially directed synthesis, e.g., a synthesis of different biopolymer species on different locations on the carrier. Methods for spatially directed biopolymer synthesis include, without limitation, light-directed synthesis, microlithography, application by inkjet, pin printing, micro channel deposition to specific locations, and sequestration with physical barriers. In general, these methods involve generating active sites on the carrier, usually by removing protecting groups and coupling to the active site a monomeric or oligomeric building block which, itself, optionally has a protected active site if further coupling of building blocks is desired.

[0018] In a preferred embodiment, the synthesis of the biopolymer species is carried out in a manner that the detectable protecting groups remain on the biopolymer species until the synthesis has been terminated. It should be noted, that the synthesis can be carried out using standard protocols.

[0019] After synthesis has been terminated, a first determination of the detectable protecting groups on the array may be carried out. By this means, qualitative and/or quantitative determination of the biopolymer species on the array is possible. The first determination step is preferably a spatially resolved determination step wherein a qualitative and/or quantitative determination of detectable protecting groups is carried out separately on different locations of the array. Techniques for spatially directed detection procedures may comprise the use of spatially resolved detectors, e.g., microscopes or detector matrices such as CCD imaging systems allowing a parallel determination on a plurality of locations on the array.

[0020] After the optional first determination, the detectable protecting groups are cleaved off. The cleavage may be carried out by known protocols according to the nature of the respective detection groups, e.g., by photochemical methods such as irradiation, or by chemical methods such as acid or base treatment. In this context it should be noted that the

detectable protection groups are preferably selected such that they are not cleaved off during the biopolymer synthesis procedure.

[0021] After the cleavage of the detectable protecting groups, a determination of the detectable protecting groups is carried out on the array in order to determine the efficacy of the protection. This determination may be qualitative and/or quantitative. If a first determination is carried out, the difference between the amount of protecting groups before and after cleavage is an indication of the amount of deprotected biopolymer and thus an indication of the deprotection rate. When no more protecting groups are detected after deprotection, the deprotection has been quantitative, otherwise a repetition of deprotection may be necessary.

[0022] In a first embodiment of the invention, the detectable protecting groups are selected from fluorescent groups, e.g., groups comprising pyrene, dansyl, stilbene, rhodamine, and/or coumarin moieties. An important feature of the fluorescent groups is that they are stable during oligonucleotide synthesis. For example, the fluorescent moiety may be combined with an acyl, e.g., a *tert*-butylphenoxyacetyl protecting group or another amino protecting group.

[0023] In a further embodiment the detectable protecting groups may be selected from radioactively detectable protecting groups. For example, radioactively detectable groups are selected from groups comprising ^{14}C, ^{32}P, and ^{3}H doped moieties which may be combined with a suitable protecting group, e.g., an amino protecting group.

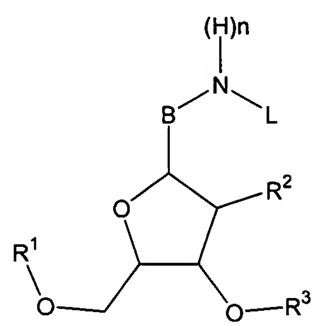
[0024] In a still further embodiment, the detectable protecting groups may be selected from electrochemically detectable protecting groups, e.g., groups comprising ferrocene and/or phenothiazine moieties. These moieties may be combined with suitable protecting groups, e.g., amino protecting groups. The determination of such protecting groups may be carried by electrochemical methods, e.g., voltammetry.

[0025] In a still further embodiment of the present invention the detectable protecting groups may be selected from UV- or IR-detectable protecting groups, e.g., from groups comprising aromatic nitro moieties, hydroxyl moieties, thiol, thioether or thiophenol

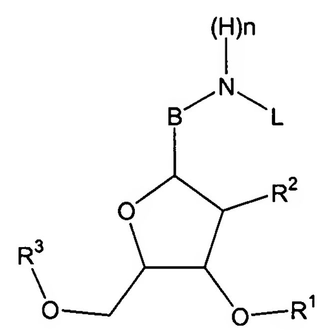
moieties, nitrile moieties, isocyanate moieties, and/or halo moieties. UV- and IR-detectable protecting groups are preferably detectable independently from the biopolymer species which has been synthesized on the array.

[0026] In a still further embodiment, the detectable protecting groups are selected from bioaffinity detectable protecting groups. Bioaffinity detectable protecting groups comprise moieties selected from partners of a bioaffinity binding pair which can be detected by their specific interaction with the respective other partner of the binding pair. Specific examples of bioaffinity detectable protecting groups are protecting groups comprising biotin, digoxin, or digoxigenin moieties. Biotin may be detected by its bioaffinity interaction with streptavidin or avidin or with anti-biotin-antibodies. Digoxin and digoxigenin may be detected by their specific interaction with respective antibodies.

[0027] In an especially preferred embodiment of the present invention the building blocks for the biopolymer synthesis are nucleotide building blocks having the general structural formulae I or II:



I



II

wherein R^1 is an hydroxy protecting group,

R^2 is -H, -(C_1-C_{10})-alkoxy, -(C_2-C_{10})-alkenyloxy, -(C_2-C_{10})-alkynyloxy, -halo, -azido, -NHR^7, -SR^7 or -OR^7, wherein R^7 is a protecting group or a reporter group,

R^3 is a phosphate, an H-phosphate or other phosphate analog group which may contain a protecting group,

B is a nucleobase or a nucleobase analog,

n is 0 or 1, and

L is a detectable protecting group, e.g., of the structure -C(O)-R, when n=1, or =CH-NR^8R, when n=0, wherein R is the residue of the reporter group and R^8 is selected from the group consisting of -(C_1-C_3)-alkyl.

[0028] In the compounds (I) or (II) the 5' or 3' hydroxy protecting group R^1 is preferably selected from optionally substituted triphenylmethyl groups, e.g., 4,4'-dimethoxy triphenylmethyl or 4-monomethoxy triphenyl, pixyl groups, photocleavable groups, e.g., p-nitrophenylpropoxy carbonyl (NPPOC) or (α-methyl)-6-nitropiperonyloxy carbonyl (MeNPOC), and substituted silyl protecting groups, e.g., *tert*-butyldimethyl silyl (TBDMS) or *tert*-butyldiphenyl silyl (TBDPS).

[0029] The group R^3 is a phosphate or phosphate analog group which may contain a protecting group, preferably a phosphitamide group, more preferably a group -$P(R^6)NR^4R^5$, wherein R^4 and R^5 are independently selected from the group consisting of -H, -(C_1-C_{10})-alkyl, -(C_2-C_{10})-alkenyl, -(C_6-C_{22})-aryl, or wherein NR^4R^5 can form together with N a 5-6-membered heterocyclic ring,
R^6 is selected from the group consisting of -(C_2-C_6)-alkenyloxy, -(C_2-C_6) alkenyl, -(C_1-C_6)-alkyl, or -(C_1-C_6)-alkoxy,
wherein each group may contain one or several substituents selected from -halo, p-nitroaryloxy and -cyano or wherein R^6 is -H.
R^6 is most preferably a 2-cyanoethyloxy group.

[0030] B is a nucleobase or nucleobase analog having at least one amino group, e.g., adenine, guanine, cytosine, or a corresponding nucleobase analog. The nucleobase or nucleobase analog is preferably capable of hydrogen bridge formation with a complementary nucleobase after incorporation into a nucleic acid molecule. Suitable nucleobase analogs are mono- and bicyclic heterocycles comprising at least one amino group wherein the heterocycle is different from natural nucleobases as described by Simons, (Advanced 2001 Chemistry Texts, Nucleoside Mimetics, Gordon and Breach Science Publishers, Amsterdam 2001, Chapter 4), for example, aza analogs of naturally occurring nucleobases, wherein a CH moiety of a purine or a pyrimidine ring is replaced

by nitrogen (such as 8-aza-adenosine) or deaza analogs of naturally occurring nucleobases wherein an N atom in the ring is substituted by a CH group (such as 7-deaza-guanosine) or combinations of aza and deaza substitutions (such as 8-aza-7-deaza-guanosine). Further preferred nucleobase analogs are C-nucleosides, e.g., deaza nucleobases wherein the N9 atom of a purine base or the N1 atom of a pyrimidine base respectively is substituted by a carbon atom (in the sp^2 configuration) such as formycine and pseudoisocytidine. Further, suitable nucleobase analogs may carry additional amino groups, e.g., 2-aminoadenosine. Further, the nucleobase or nucleobase analog may be substituted, wherein the substituent, e.g., a C_1-C_3 alkyl or alkoxy, a C_2-C_3 alkenyl or alkenyloxy, a C_2-C_3 alkynyl or alkynyloxy, and/or a halo substituent, is preferably compatible with the formation of hydrogen bridges to a complementary base. Preferred positions for substituents are C5 in pyrimidine bases, C8 in purine bases, and C7 in deaza purine bases.

[0031] Furthermore, the present invention relates to a nucleic acid synthesis building block having the general structural formulae I or II as described above. The building blocks are suitable for the production of nucleic acid arrays.

[0032] Furthermore, the invention relates to a reagent kit for the synthesis of nucleic acid arrays comprising at least one nucleic acid synthesis building block as described above. For example, the reagent kit may comprise at least two different nucleic acid synthesis building blocks each carrying different detectable protecting groups. The different detectable protecting groups are preferably detectable independent from each other. For example, the reagent kit may comprise at least two different independently detectable fluorescent protecting groups or combinations of different embodiments of detectable protecting groups, e.g., fluorescent and bioaffinity detectable protecting groups or other combinations of different protecting groups.

Specific Embodiments

Example. Synthesis of a stilbene-labelled nucleic acid synthesis building block

[0033] Starting from 4-hydroxystilbene the desired stilbene-labeled building block was synthesized in the following steps A) to F) according to the scheme in Figure 5:

*A)Synthesis of (4-styryl-phenoxy)-acetic acid ethyl ester **(a)***

[0034] 550 mg sodium hydride (60% suspension in paraffin oil) were added to a solution of 2 g trans 4 hydroxystilbene in 70 ml dry dioxane. After stirring for 1 h at room temperature 1.2 ml of iodoacetic acid ethylester were added. The mixture was refluxed for 1 h and then cooled to room temperature. After filtration the solvent was removed by using a rotary evaporator and the remainder was purified by column chromatography on silica gel with hexane/acetic acid ethyl ester 2:1 as eluent. Yield: 1,0g.

*B) Synthesis of (4-styryl-phenoxy)-acetic acid **(b)***

[0035] 20 ml 10 Mol NaOH were added within 20 min to a solution of 1 g (4-styryl-phenoxy) acetic acid ethyl ester (a) in 60 ml dioxane. After stirring for 5 h the pH was adjusted to pH 2 with 10 Mol HCl. The mixture was extracted with acetic acid ethyl ester. The organic layer was separated, washed with water and dried over sodium sulfate. The solvent was evaporated by using a rotary evaporator. The remainder was dried in vacuum for 5 days with calcium chloride as desiccant. Yield: 0.81 g.

*C) Synthesis of N6(4-styryl-phenoxy)-acetyl-4.2(-)-3',5'-O-(1,1,3,3-tetraisopropyl-1,3-disiloxanediyl) adenosine **(c)***

[0036] 0.81 g (4-styryl-phenoxy)-acetic acid (b) was dissolved in a mixture from 15 ml dioxane and 3 ml methylene chloride. After adding 2.6 ml oxalyl chloride at 4-8°C the mixture was stirred at room temperature for 4 h. The solvents were removed by using a rotary evaporator. The remainder was dissolved in 10 ml methylene chloride. This solution was dropped at room temperature to a solution of 1.35 g 4.2 (-)-3',5'-O-(1,1,3,3-tetraisopropyl-1,3-disiloxanediyl) adenosine in 30 ml methylene chloride and 3 ml pyridine. The mixture was then stirred overnight. The solvent was removed by using a rotary evaporator and the remainder was purified by column chromatography on silica gel with toluene/acetic acid ethyl ester/methanol 4:1:1. Yield: 1.5 g.

*D) Synthesis of N6 (4-styryl-phenoxy)-acetyl adenosine **(d)***

[0037] 1.5 g of N6(4-styryl-phenoxy)-acetyl- 4.2 (-)-3',5'-O-(1.1.3.3-tetraisopropyl-1,3-disiloxanediyl) adenosine (c) were dissolved in 90 ml 1 M tetrabutylammoniumfluoride THF solution and stirred overnight. The solvent was removed by using a rotary evaporator and the remainder was purified by column chromatography on silica gel with acetic acid ethyl ester/methanol 2:1. Yield: 0.56 g.

*E) Synthesis of 5'-(p,p'-dimethoxytrityl)- N6 (4-styryl-phenoxy)-acetyl adenosine **(e)***

[0038] Within 1 h a solution of 0.46 g dimethoxytritylchloride in 10 ml dry pyridine was dropped at room temperature to a solution of 0.56 g of N6 (4-styryl-phenoxy)-acetyl adenosine **(d)** in 10 ml dry pyridine. After stirring overnight the solvent was evaporated using a rotary evaporator. The residue was dissolved in 300 ml acetic acid ethyl ester and was washed with 100 ml water. The organic phase was separated and dried with sodium sulfate. After filtration the solvent was removed by using a rotary evaporator.

[0039] Purification was performed by column chromatography on silica gel. Therefore the crude product was dissolved in a methylene chloride/methanol 10:1 mixture containing 0.1% triethylamine. The solution was applied on an l= 40 cm/d=6.9 cm column filled with silica gel 60 (0.063-0.200 mm). The product was eluted with a methylene chloride/methanol 10:1 mixture containing 0.1% triethylamine. TLC (Silicagel): methylene chloride/methanol 10:1 mixture containing 0.1% triethylamine: Rf: 0.56. Yield: 200 mg

*F) Synthesis of 2-Cyanoethylphosphoramidite of 5'-(p,p'-dimethoxytrityl)- N6 (4-styryl-phenoxy)-acetyl adenosine **(f)***

[0040] At room temperature under argon 0.56 μL of N ethyldiisopropylamine and then 73 μl chloro-2 cyanoethoxy diisopropylaminophosphane were added to a solution of 200 mg of 5'-(p,p'-dimethoxytrityl)- N6 (4-styryl-phenoxy)-acetyl adenosine **(e)** in 5 ml methylene chloride. The solvent was removed by using a rotary evaporator. The residue was dissolved in 50 ml acetic acid ethyl ester and washed two times with 10 ml 5%

aqueous sodium hydrogen carbonate solution. The organic phase was separated and dried with sodium sulfate. After filtration the solvent was removed by using a rotary evaporator.

[0041] Purification was performed by column chromatography on silica gel. Therefore the crude product was dissolved in a methylene chloride/acetone/triethylamine 45:5:0.5 mixture. The solution was applied on a l=32 cm/d=4.5 cm column filled with silica gel 60 (0.063-0.200 mm). The product was eluted with an acetic acid ethyl ester/hexane 1:1 mixture containing 0.1% triethylamine. TLC (Silica gel, methylene chloride/acetone/triethylamine 45:5:0.5): Rf: 0.70. NMR(d6-DMSO) 1H(300 MHz): 10.97 s[1H], 8.64 s[1H], 8.56s [1H], 6.78-7.5 m [24H], 6.50 s, broad [1H], 5.20 s[2H], 4.82 s, broad [1H], 4.16 s, broad [1H], 3.64 s [6H], 3.55 s, broad [2H], 3.64 s, broad [2H], 3.23 s, broad [2H], 3.11 m, broad [1H], 2.73 dt[2H], 2.55 m, broad [1H] 1.22 dd [12H] 31P(300 MHz): 148.89 d. Yield: 120 mg.